

Speech Processing 2015/16

1st Test

March 17th 2016

Please identify this form with your name and student number in the reserved spaces in the two sheets. Wrong answers to True/False questions will be penalized. The phonetic symbols should use the SAMPA alphabet. Portuguese speaking students should use the alphabet for the Lisbon accent. Foreign students may use the SAMPA alphabet for English.

Name:	
Number:	
Group number:	
Place of birth:	

1. Classify as True (T) or False (F) on the left side of each item.

- (a) AMDF allows pitch detection without having to compute multiplications.
- (b) Error minimization in linear prediction favors the valleys between formants.
- (c) The cepstrum of a decaying exponential a^n decays twice as rapidly as the exponential.
- (d) The outer ear protects against persistent strong sounds.
- (e) The coincidence of F0 with F1 occurs more often for male speech than for female speech (take into account the range of F1 values that were obtained in the Lab for all vowels).

2. Fill in the blanks:

The cut-off frequency of a _____, with high-pass characteristics, is proportional to the length of the cavity _____ the constriction.

3. Write the Z-transform for the inverse filter used in linear prediction.

--

4. Which equivalent representations of linear prediction coefficients have a limited range (for a stable filter)?

5. How do the following coefficients reproduce the intensity / loudness conversion?

- (a) MFCC _____
- (b) PLP _____

6. A sentence has been recorded with 16kHz, 16 bits per sample, duration 2.5s. 13 MFCCs were computed per frame, with windows 30 ms long and interval 10 ms. Their values were used to train a classifier that yielded 25% silence frames, 60% voiced frames and 15% unvoiced frames. (Please indicate all the steps you computed in answering the following questions).

(a) What is the total size of the recorded file in kilobytes?

(b) How many frames correspond to speech?

(c) How many MFCC values were computed in total?

7. Consider the following sentence in European Portuguese:

O candidato (...) p̃ĩç ½s em causa a seriedade dos advers̃ĩç ½rios neste processo eleitoral.

If you are not a native Portuguese speaker, you may use the Google translation in English, and the SAMPA alphabet for that language:

The applicant (...) questioned the seriousness of opponents in this electoral process.

(a) Write the broad phonetic transcription.

(b) Complete the sentence with the name of a fictitious candidate so that the sentence contains at least one example of each palatal sound in EP (orthography and broad phonetic transcription).

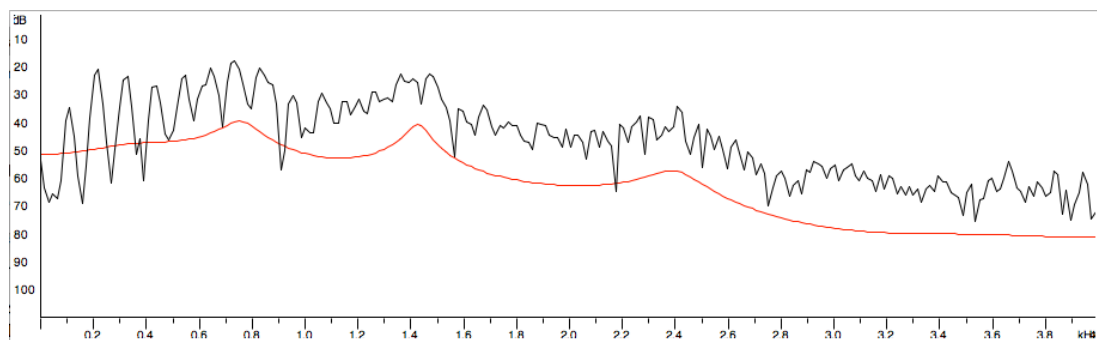


Figure 1

8. Figure 1 presents the short-time Fourier transform (magnitude) of a vowel segment, including its LPC spectral envelope.

(a) Indicate approximate values for F0, F1, F2 and F3 (Hz).

F0= _____ F1= _____

F2 = _____ F3 = _____

(b) To which of the 3 vowels of the extremes of the vowel triangle does it correspond? _____

(c) May it correspond to a male voice? _____

(d) Which of the three formants has poles closest to the circumference (unit radius) ? _____

9. For the magnitude spectrum shown in Figure 1, compute the angles corresponding to the first three formants (in degrees) in the complex plane.

10. Which is the minimum size of the FFT window (power of 2, in points) to obtain the magnitude spectrum shown in Figure 1? Justify briefly including computations.

11. Identify the 10-digit sequence by inspecting the spectrograms and waveforms in the next page. There are no repeated digits. The recordings correspond to telephone speech.

--	--	--	--	--	--	--	--	--	--

12. Could the magnitude spectrum in Figure 1 be obtained from the recordings used for the digit spectrograms in the next page? If yes, indicate the approximate location(s) in the time axis. If no, justify briefly.

1	1.5
2	1.0
3	1.4
4	1.5
5	1.0
6	1.5
7	2.6
8	2.5
9	1.5
10	1.0
11	3.0
12	1.5

Name:	
Number:	
Group number:	
Place of birth:	

